

# The Road to DocuComp

Ted Habermann, NOAA/NESDIS/NGDC

NCDC Metadata Workshop

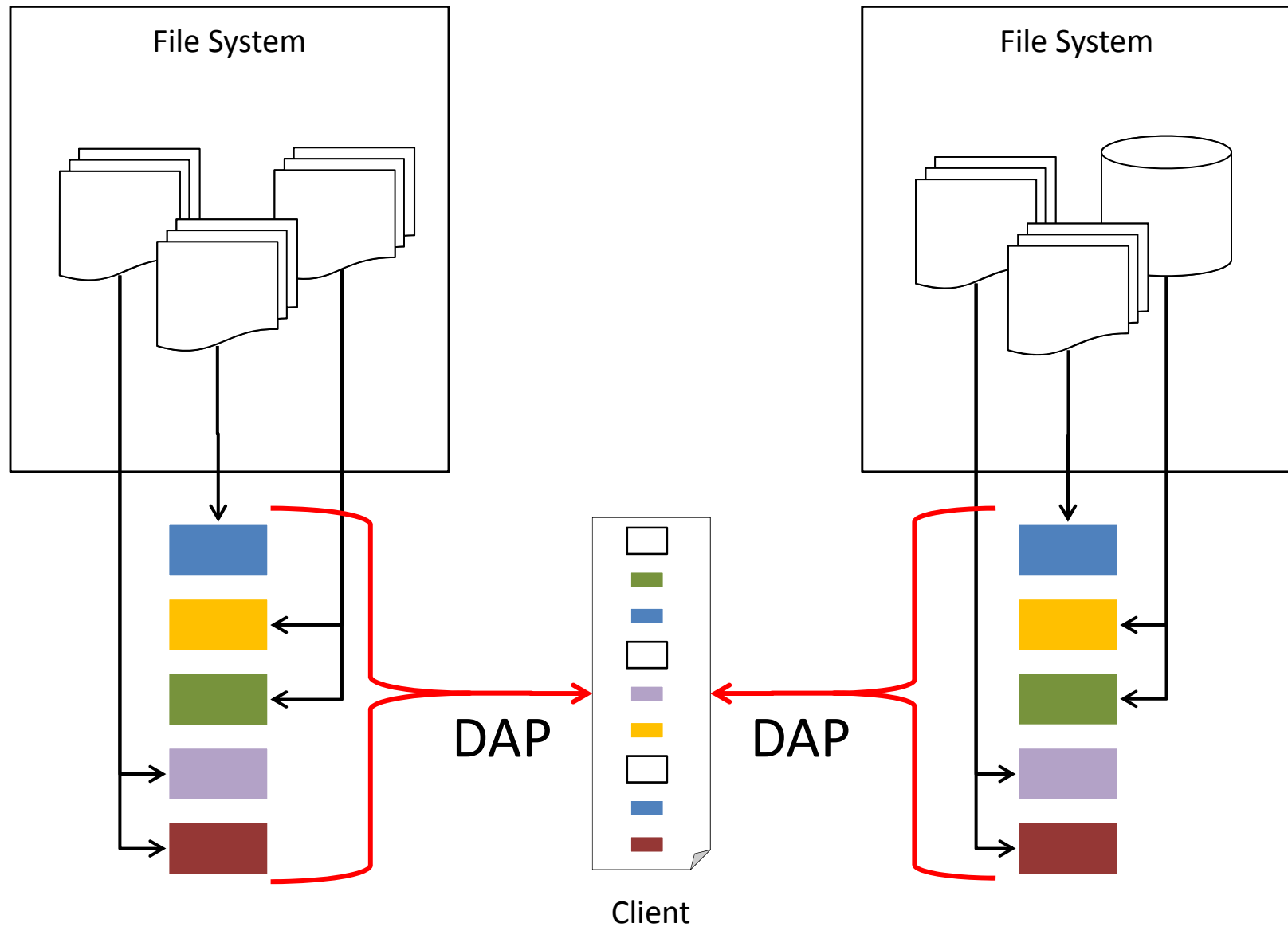


Database normalization is the process of organizing the fields and tables of a relational database to minimize redundancy and dependency.

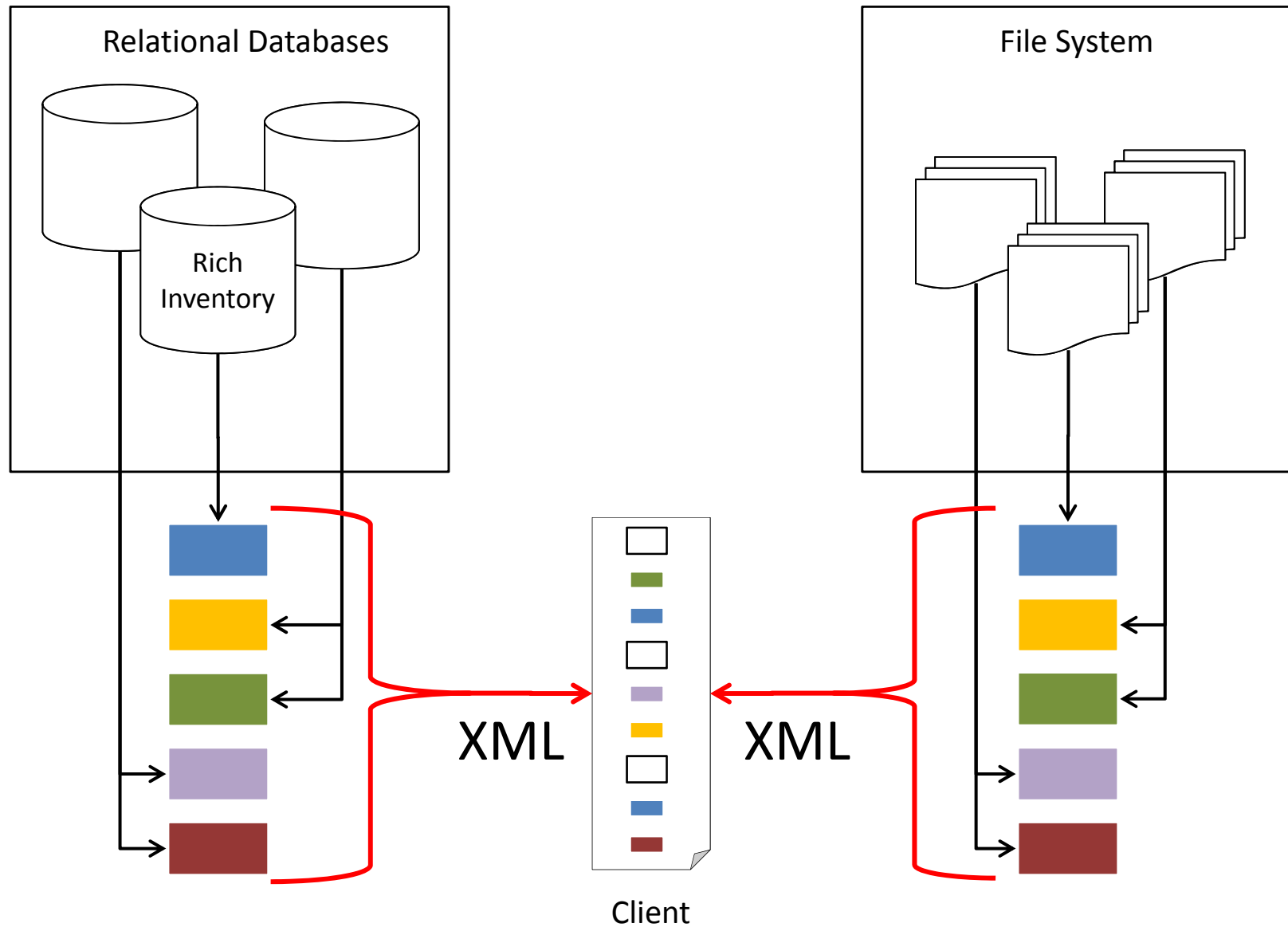
The objective is to isolate data so that additions, deletions, and modifications can be made once and then propagated through the rest of the database via defined relationships.

DocuComp brings the benefits of normalization to metadata using reusable components and xlink instead of tables and foreign keys.

# Persistence vs. Transport – OPeNDAP



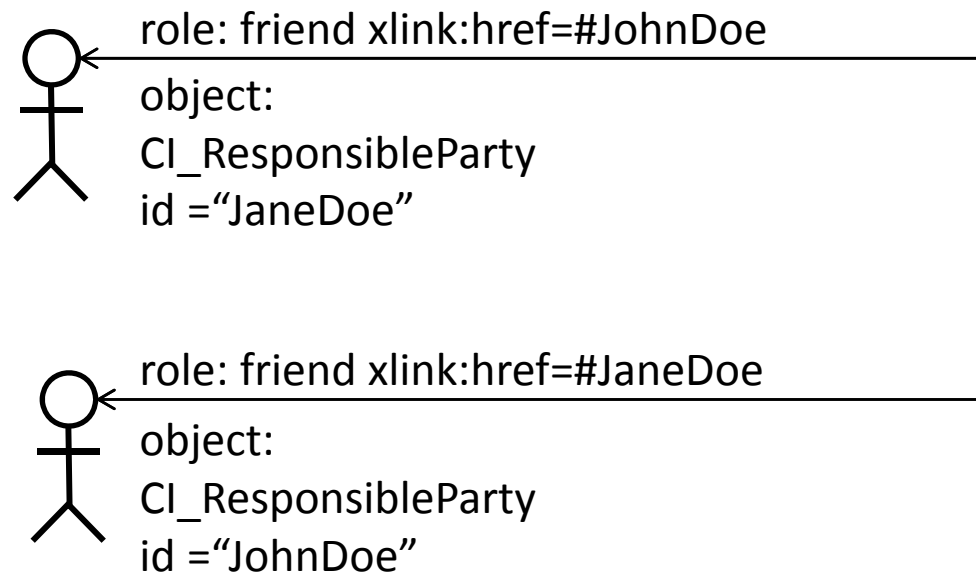
# Persistence vs. Transport - Documentation



# Roles, Identifiers and References

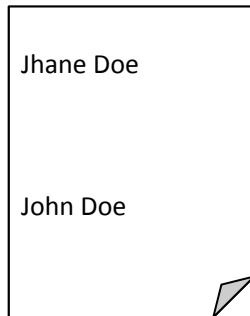
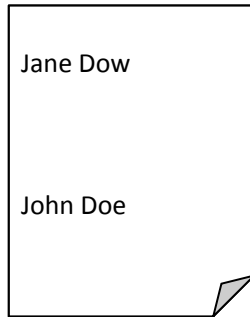
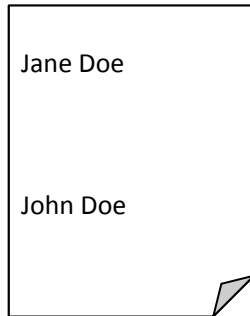
ID	Name	Friend_ID
JaneDoe	Jane Doe	JohnDoe
MaryBrown	Mary Brown	TomBrown
TomBrown	Tom Brown	MaryBrown
JohnDoe	John Doe	JaneDoe

This table has four people, all of which have identifiers (ID), names, and friends. The friends are identified by foreign keys (Friend\_ID).



This information is represented in XML as objects with ids and references with xlink.

# Jane or Jhane?



```
<gmd:contact>
  <gmd:CI_ResponsibleParty>
    <gmd:individualName>
      <gco:CharacterString>Jhane Doe</gco:CharacterString>
    </gmd:individualName>
    <gmd:organisationName>
      <gco:CharacterString>
        DOC/NOAA/NESDIS/NGDC > National Geophysical Data Center, NESDIS, NOAA, U.S. Department of Commerce
      </gco:CharacterString>
    </gmd:organisationName>
    <gmd:contactInfo>
      <gmd:CI_Contact>
        <gmd:address>
          <gmd:CI_Address>
            <gmd:electronicMailAddress>
              <gco:CharacterString>ngdc.info@noaa.gov</gco:CharacterString>
            </gmd:electronicMailAddress>
          </gmd:CI_Address>
        </gmd:address>
        <gmd:onlineResource>
          <gmd:CI_OnlineResource>
            <gmd:linkage>
              <gmd:URL>http://www.ngdc.noaa.gov</gmd:URL>
            </gmd:linkage>
          </gmd:CI_OnlineResource>
        </gmd:onlineResource>
      </gmd:CI_Contact>
    </gmd:contactInfo>
    <gmd:role>
      <gmd:CI_RoleCode codeListValue="pointOfContact">pointOfContact</gmd:CI_RoleCode>
    </gmd:role>
  </gmd:CI_ResponsibleParty>
</gmd:contact>
```

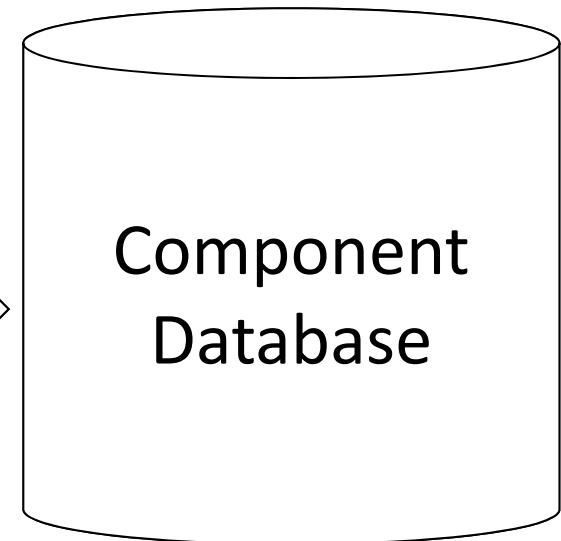
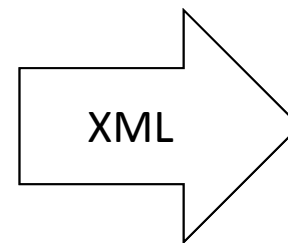
# Jane as Reusable Component

```
<gmd:contact>
  <gmd:CI_ResponsibleParty UUID="19dcdfc0-dd85-11e1-9b23-0800200c9a66">
    <gmd:individualName>
      <gco:CharacterString>Jane Doe</gco:CharacterString>
    </gmd:individualName>
    <gmd:organisationName>
      <gco:CharacterString>
        DOC/NOAA/NESDIS/NGDC > National Geophysical Data Center, NESDIS, NOAA, U.S. Department of Commerce
      </gco:CharacterString>
    </gmd:organisationName>
    <gmd:contactInfo>
      <gmd:CI_Contact>
        <gmd:address>
          <gmd:CI_Address>
            <gmd:electronicMailAddress>
              <gco:CharacterString>ngdc.info@noaa.gov</gco:CharacterString>
            </gmd:electronicMailAddress>
          </gmd:CI_Address>
        </gmd:address>
        <gmd:onlineResource>
          <gmd:CI_OnlineResource>
            <gmd:linkage>
              <gmd:URL>http://www.ngdc.noaa.gov/mgg/</gmd:URL>
            </gmd:linkage>
          </gmd:CI_OnlineResource>
        </gmd:onlineResource>
      </gmd:CI_Contact>
    </gmd:contactInfo>
    <gmd:role>
      <gmd:CI_RoleCode codeListValue="pointOfContact">pointOfContact</gmd:CI_RoleCode>
    </gmd:role>
  </gmd:CI_ResponsibleParty>
</gmd:contact>
```

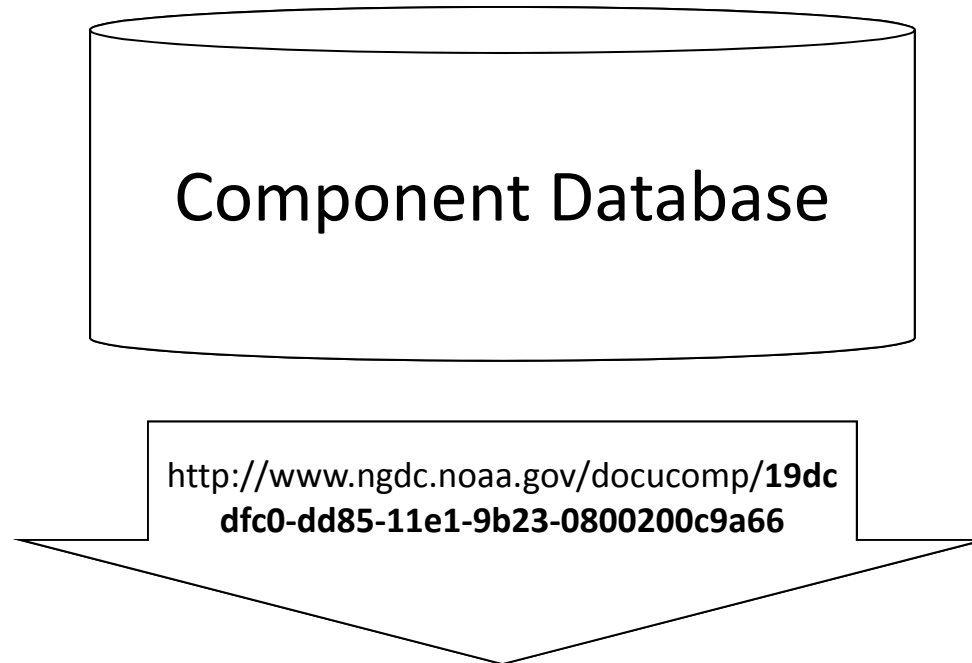
object:

CI\_ResponsibleParty

id = Universally Unique ID (UUID)

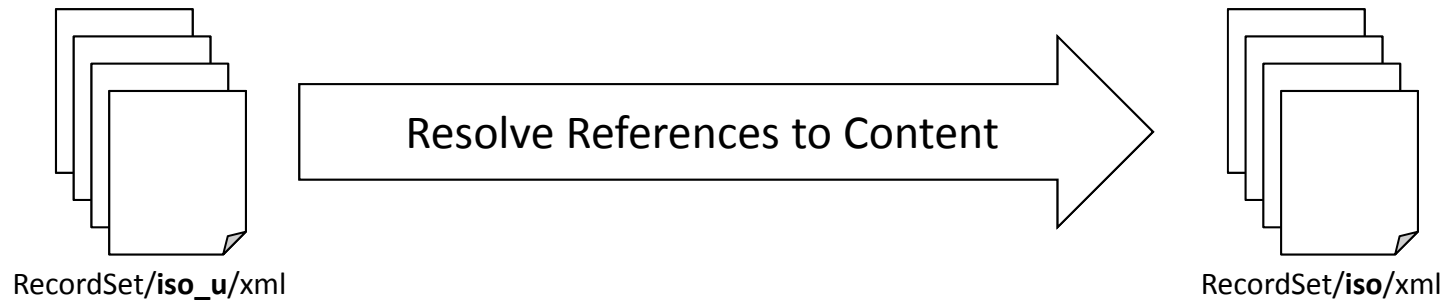


# Docucomp = Database + Restful Service



```
<gmd:CI_ResponsibleParty UUID="19dcdfc0-dd85-11e1-9b23-0800200c9a66">  
  <gmd:individualName>  
    <gco:CharacterString>Jane Doe</gco:CharacterString>  
  </gmd:individualName>  
  ...  
</gmd:CI_ResponsibleParty>
```

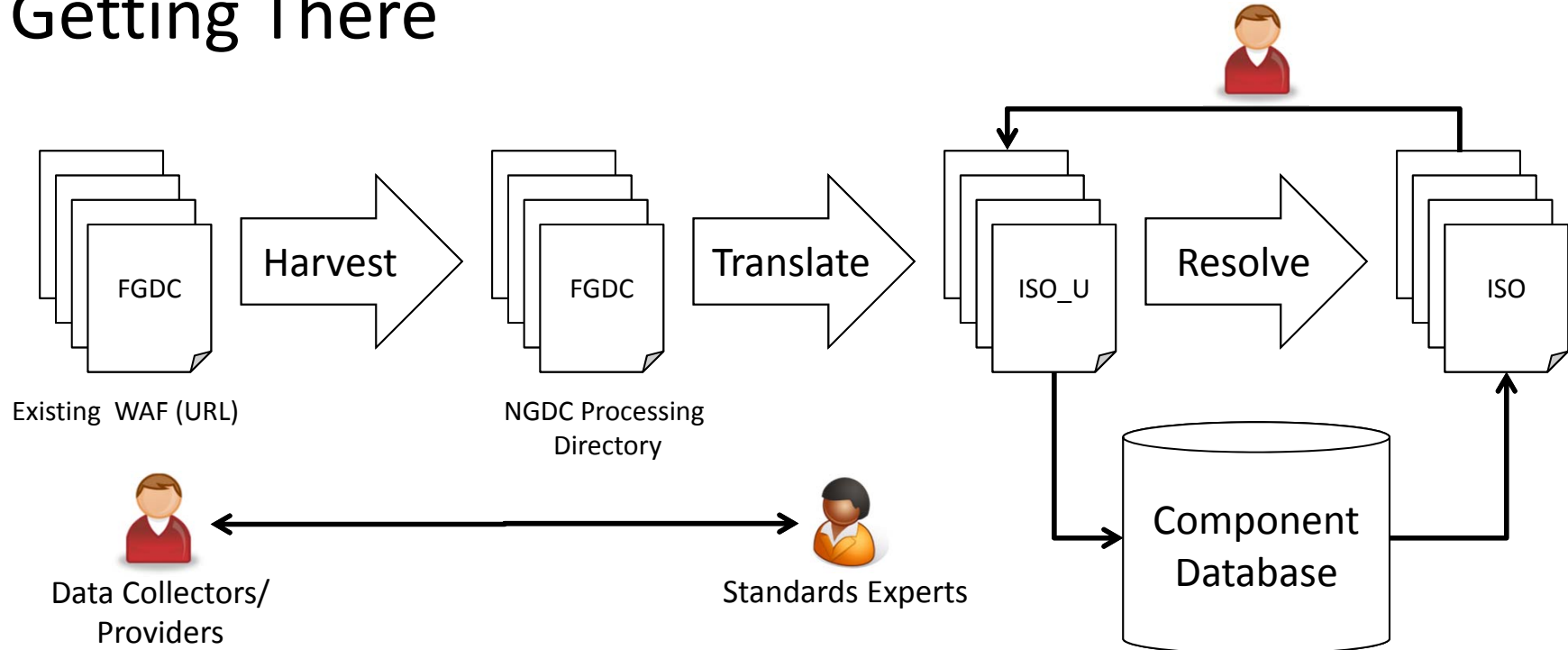
# Component Resolution



```
<gmd:contact xlink:href="http://www.ngdc.noaa.gov/docucomp/19dcdfc0-dd85-11e1-9b23-0800200c9a66" xlink:title="Jane Doe"/>
<gmd:CI_ResponsibleParty UUID="19dcdfc0-dd85-11e1-9b23-0800200c9a66">
  <gmd:individualName>
    <gco:CharacterString>Jane Doe </gco:CharacterString>
  </gmd:individualName>
  <gmd:organisationName>
    <gco:CharacterString>
      DOC/NOAA/NESDIS/NGDC > National Geophysical Data Center, NESDIS, NOAA, U.S. Department of Commerce
    </gco:CharacterString>
  </gmd:organisationName>
  <gmd:contactInfo>
    ...
  </gmd:contactInfo>
  <gmd:role>
    <gmd:CI_RoleCode codeListValue="pointOfContact">pointOfContact</gmd:CI_RoleCode>
  </gmd:role>
</gmd:CI_ResponsibleParty>
</gmd:contact>
```



# Getting There



## Iterations:

Translation Proofing: identify un-translated content and fix

Rubrics: evaluate completeness and improvement opportunities

Validation: schematron and schema validation

Link checking: check all online resources

Multiple views: Translation of ISO to multiple views

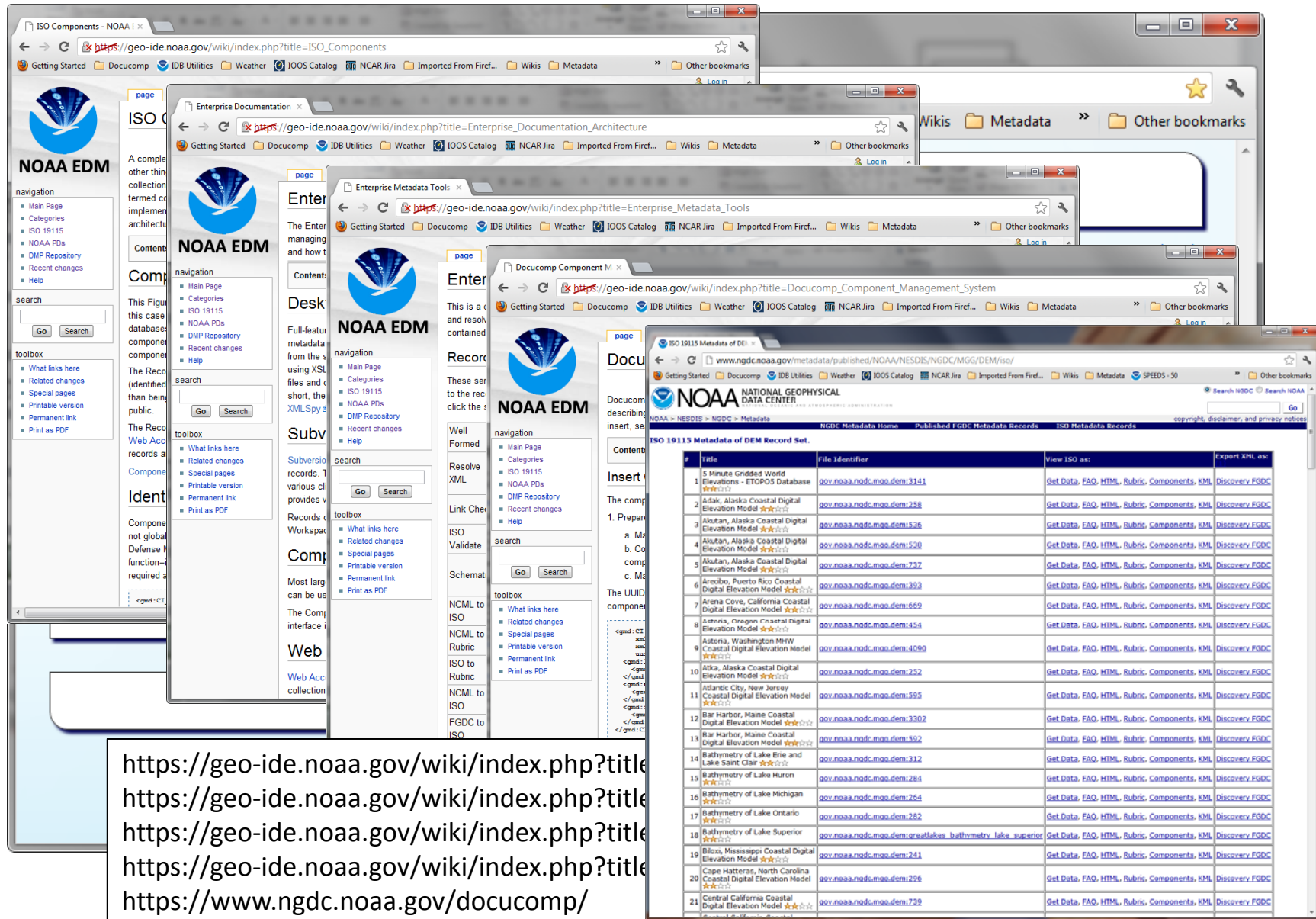
Consistency Checker: find repeated content and candidate components

DocuComp: create and store components

Metadata editor: add references into metadata records



# Resources



The collage displays several web browser windows from the NOAA and NGDC websites. The windows are overlapping, showing different parts of the site's structure and content.

The windows visible include:

- ISO Components - NOAA**: A page with a navigation menu and a search bar.
- Enterprise Documentation Architecture**: A page with a navigation menu and a search bar.
- Enterprise Metadata Tools**: A page with a navigation menu and a search bar.
- Docucomp Component Management System**: A page with a navigation menu and a search bar.
- ISO 19115 Metadata of DEM**: A page showing a table of metadata records for DEM (Digital Elevation Model) data.

The **ISO 19115 Metadata of DEM** table contains the following data:

#	Title	File Identifier	View ISO as:	Export XML as:
1	5 Minute Gridded World Elevations - ETOPOS Database	gov.noaa.nodc.moa.dem:3141	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
2	Adak, Alaska Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:238	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
3	Adak, Alaska Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:536	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
4	Adak, Alaska Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:538	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
5	Adak, Alaska Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:737	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
6	Arecibo, Puerto Rico Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:393	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
7	Arena Cove, California Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:669	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
8	Astoria, Oregon Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:434	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
9	Astoria, Washington MHW Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:4090	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
10	Atka, Alaska Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:252	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
11	Atlantic City, New Jersey Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:595	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
12	Bar Harbor, Maine Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:3302	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
13	Bar Harbor, Maine Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:392	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
14	Bathymetry of Lake Erie and Lake Saint Clair	gov.noaa.nodc.moa.dem:312	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
15	Bathymetry of Lake Huron	gov.noaa.nodc.moa.dem:284	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
16	Bathymetry of Lake Michigan	gov.noaa.nodc.moa.dem:264	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
17	Bathymetry of Lake Ontario	gov.noaa.nodc.moa.dem:282	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
18	Bathymetry of Lake Superior	gov.noaa.nodc.moa.dem:greatlakes_bathymetry_lake_superior	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
19	Blount, Mississippi Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:241	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
20	Cape Hatteras, North Carolina Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:296	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC
21	Central California Coastal Digital Elevation Model	gov.noaa.nodc.moa.dem:739	Get Data, FAQ, HTML, Rubric, Components, KML	Discovery FGDC

Below the collage, there are five lines of URLs:

- [https://geo-ide.noaa.gov/wiki/index.php?title=ISO\\_Components](https://geo-ide.noaa.gov/wiki/index.php?title=ISO_Components)
- [https://geo-ide.noaa.gov/wiki/index.php?title=Enterprise\\_Documentation\\_Architecture](https://geo-ide.noaa.gov/wiki/index.php?title=Enterprise_Documentation_Architecture)
- [https://geo-ide.noaa.gov/wiki/index.php?title=Enterprise\\_Metadata\\_Tools](https://geo-ide.noaa.gov/wiki/index.php?title=Enterprise_Metadata_Tools)
- [https://geo-ide.noaa.gov/wiki/index.php?title=Docucomp\\_Component\\_Management\\_System](https://geo-ide.noaa.gov/wiki/index.php?title=Docucomp_Component_Management_System)
- <https://www.ngdc.noaa.gov/docucomp/>

Questions?

